# *SECTION II: FREE RESPONSE*
## *Part A*

**90 minutes**
**Suggested time: 65 minutes**

1. A shipment of computer chips consists of 10,000 units. The manufacturer claims that the probability of selecting a defective unit is 0.003.
   a. How many defective units should the receiver of the shipment expect?

   b. What is the probability that the receiver of the shipment will get more defective units than expected?

   c. What is the probability that fewer than 18 units will be defective in this shipment?

2. A manufacturer has created a pole from a new material and believes that pole-vaulters can improve the heights of their vaults by using this new type of pole.
   a. In order to test this hypothesis, the manufacturer visits Kennett High School. He selects a student who is new to vaulting. He records a sample of 25 vaults where the student uses a pole made from the customary material. Then he records a sample of 25 vaults where the student uses a pole constructed from the new material. He analyzes the data and finds a significant improvement in heights with the new pole. Comment on the design of this experiment.

b. Could you improve upon the design of the experiment in part (a)? If you answer yes, describe *your* design.

c. As pole-vaulters improve their performances, longer poles become a necessity. Describe an experiment to test the new material that takes different pole lengths into account.

3. Suppose that the weights of a name-brand cereal vary normally with mean $\mu = 11.13$ oz and standard deviation $\sigma = 0.08$ oz. The advertised weight is 11 oz. For the equivalent generic brand with an advertised weight of 11 oz, the weights vary normally with mean $\mu = 11.15$ oz and standard deviation $\sigma = 0.16$ oz.

   a. For each of the brands, find the probability that the weight of a box of cereal will be less than the advertised weight.

   b. For each of the brands, find the probability that the average weight for a purchase of four boxes of cereal will exceed 11.25 oz.

   c. If you believe strongly in getting what you pay for, that is, you want to make sure you get at least the weight advertised on the box, which brand of cereal are you more likely to buy and why?

   d. If you wanted to get the most for your money, that is, you would like to get much more than the advertised weight, which brand of cereal are you more likely to buy and why?

4. Based on results from a random sample of 35 individuals, a company advertises that individuals who use its diet supplement lose an average of 4.6 lb with a standard deviation of 1.2 lb during the first week of dieting.

   a. A graph of the data reveals no strong skew or outliers. Does this provide evidence of a significant weight loss?

   b. The advertisement fails to mention that these results were part of a double-blind experiment to compare the supplement with a placebo. In the control group, the weight loss was 3.7 lb with a standard deviation of 2.3 lb in the first week of dieting. The control group consisted of a random sample of 32 individuals. A graph of the data for the control group reveals no strong skew or outliers. Is there evidence to show that the group taking the diet supplement lost more weight than the control group in the first week?

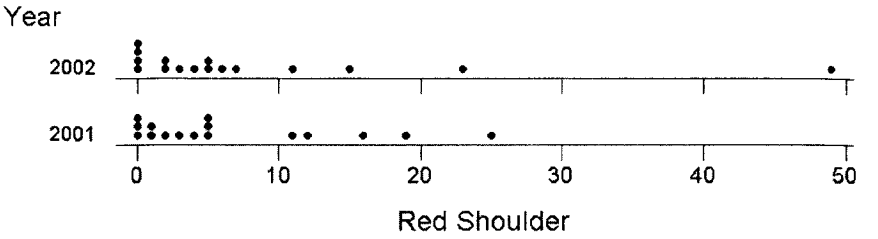
5. In eastern Pennsylvania, the peak migration for red-shouldered hawks is October 15–31. During 2001 and 2002, the owner of a mountaintop home recorded the number of these hawks that migrated past her home each day during the peak migration season. The following summary statistics resulted:

| Variable | Year | N | Mean | Median | TrMean | StDev |
|---|---|---|---|---|---|---|
| Red Shoulder | 2001 | 16 | 6.81 | 4.50 | 6.00 | 7.62 |
| | 2002 | 16 | 8.25 | 4.50 | 5.93 | 12.53 |

| Variable | Year | SE Mean | Minimum | Maximum | Q1 | Q3 |
|---|---|---|---|---|---|---|
| Red Shoulder | 2001 | 1.90 | 0.00 | 25.00 | 1.00 | 11.75 |
| | 2002 | 3.13 | 0.00 | 49.00 | 0.50 | 10.00 |

   a. Are there any outliers in the distribution of migration data for red-shouldered hawks for 2002? Describe the procedure you used to determine outliers, and justify your answer based on that procedure.

b. Dotplots for the distributions of migration data for both years are given.




If you said there were outliers in part (a), would the value(s) still be outliers if we combined the results from 2001 and 2002? What is the total number of outliers for the combined data?

c. Write a few sentences comparing the migration of red-shouldered hawks in 2001 and 2002.

d. Using the combined results from 2001 and 2002, if a friend visits during the peak migration season, what is the probability the friend will see on a given day 20 or more red-shouldered hawks?

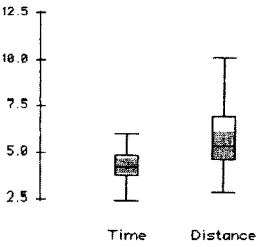
e. Using the combined results from 2001 and 2002, what is the most probable number of red-shouldered hawks the same friend would see on a given day during the peak migration season?

# *Part B*

**Suggested time: 25 minutes**

6. A hiker records distances (in miles) and times (in hours) for a random sample of 25 of her hikes, resulting in the following information.

| Summary of | Distance | Summary of | Time |
|---|---|---|---|
| No Selector | | No Selector | |
| **Count** | 25 | **Count** | 25 |
| **Mean** | 5.952 | **Mean** | 3.7 |
| **Median** | 5.7 | **Median** | 3.5 |
| **StdDev** | 2.71356 | **StdDev** | 1.54448 |
| **IntQRange** | 2.975 | **IntQRange** | 2.0625 |



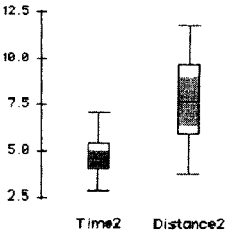

Dependent variable is: **Time**
No Selector
R squared = 77.3% R squared (adjusted) = 76.3%
s = 0.7515 with 25 - 2 = 23 degrees of freedom

| **Source** | **Sum of Squares** | **df** | **Mean Square** | **F-ratio** |
|---|---|---|---|---|
| Regression | 44.2594 | 1 | 44.2594 | 78.4 |
| Residual | 12.9906 | 23 | 0.564807 | |

| **Variable** | **Coefficient** | **s.e. of Coeff** | **t-ratio** | **prob** |
|---|---|---|---|---|
| Constant | 0.721346 | 0.3685 | 1.96 | 0.0625 |
| Distance | 0.500446 | 0.05653 | 8.85 | ≤ 0.0001 |

A friend also records distances (in miles) and times (in hours) for a random sample of 20 of his hikes, resulting in the following information.

| Summary of | Distance2 | Summary of | Time2 |
|---|---|---|---|
| No Selector | | No Selector | |
| **Count** | 20 | **Count** | 20 |
| **Mean** | 7.585 | **Mean** | 4.5875 |
| **Median** | 7.65 | **Median** | 4.5 |
| **StdDev** | 2.4577 | **StdDev** | 1.2677 |
| **IntQRange** | 3.7 | **IntQRange** | 1.375 |




Dependent variable is: **Time2**
No Selector
R squared = 85.5% R squared (adjusted) = 84.7%
s = 0.4952 with 20 - 2 = 18 degrees of freedom

| **Source** | **Sum of Squares** | **df** | **Mean Square** | **F-ratio** |
|---|---|---|---|---|
| Regression | 26.1202 | 1 | 26.1202 | 107 |
| Residual | 4.41416 | 18 | 0.245231 | |

| **Variable** | **Coefficient** | **s.e. of Coeff** | **t-ratio** | **prob** |
|---|---|---|---|---|
| Constant | 0.968919 | 0.3677 | 2.64 | 0.0168 |
| Distance2 | 0.477071 | 0.04623 | 10.3 | ≤ 0.0001 |

These friends are trying to determine whether or not they would be well suited to hike together.

a. One of the friends believes that hiking distances are important to determining compatibility. Is there any evidence of a significant difference in hiking distances between the two hikers?

b. The other friend believes it is the length of the hike in time that best determines compatibility. Is there any evidence of a significant difference in hiking times between the two hikers?

c. Should the two friends hike together? Use your results from parts (a) and (b) to support your answer.

d. Rather than look at the choice of hikes, which would be reflected in the distances and times of the hikes reported, a mutual friend suggests that the two should instead look at their hiking rate. This friend conjectures that hikers with rates of speed within 0.1 mph of one another make compatible hikers. Record the speed at which these friends hike. Would your answer to part (c) change if their mutual friend is correct in his conjecture?